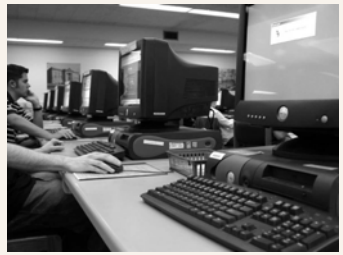




The New Academic Library:

Building Repositories to Support Changing Scholarly and Research Processes

Tyler O. Walters
Associate Director, Technology and Resource Services
Library and Information Center
Georgia Institute of Technology

ACRL Baltimore, March 30-31, 2007

Georgia Tech Library and Information Center

The World of Repositories

- What will academic libraries look like in the future?
  - What information will they hold?
  - What resources will be accessed?
  - What services will they provide?
- Growing # of libraries are collecting, organizing, disseminating intellectual output of parent institutions
- Education and research programs -- voluminous and diverse and needs to be managed well for long-term
- Institutional repositories (IRs), are proliferating at a great rate (organize, access, begin preservation)
- 843 IRs in OpenDOAR… in last 5 yrs!

Georgia Tech Library and Information Center

Repository Content

- Libraries focused on two things:
  - users and their collections
- To better understand why we need repositories and how they should function…

  …focus on contemporary processes of academic research, learning, and communication
- Examine information outputs to determine how to they should function in digital object mgmt. framework
- Only then can we identify tangible intellectual outputs and create / maintain comprehensive colls.
- Without well-designed functions, IR of little use to users

College / university's digital intellectual output includes diverse items, such as:

- Annual reports
- Computer programs
- Conference papers
- Data sets
- Learning/complex objects (digitally captured courses, multimedia simulations/visualizations, captured notes of faculty/students, etc.)
- Lecture series materials
- Models
- Pre-prints/post-prints
- Proceedings
- Research reports
- Simulations/visualizations
- Technical reports and working papers
- Web pages
- White papers

Georgia Tech Library and Information Center

Four Types of Intellectual Output

- The four major categories of output being reviewed are:

  1. Faculty and researchers' scholarly communications
     - (i.e., pre-/post-prints, journal articles, conference papers, research reports, technical papers, etc.)
  2. Student intellectual output
  3. Learning objects and other multimedia-based works
  4. Digital research data sets

- Resources pose new challenges, also present new opportunities
- Libraries extending capabilities to manage these resources for future
  - becoming integral to academic knowledge dissemination processes

Georgia Tech Library and Information Center

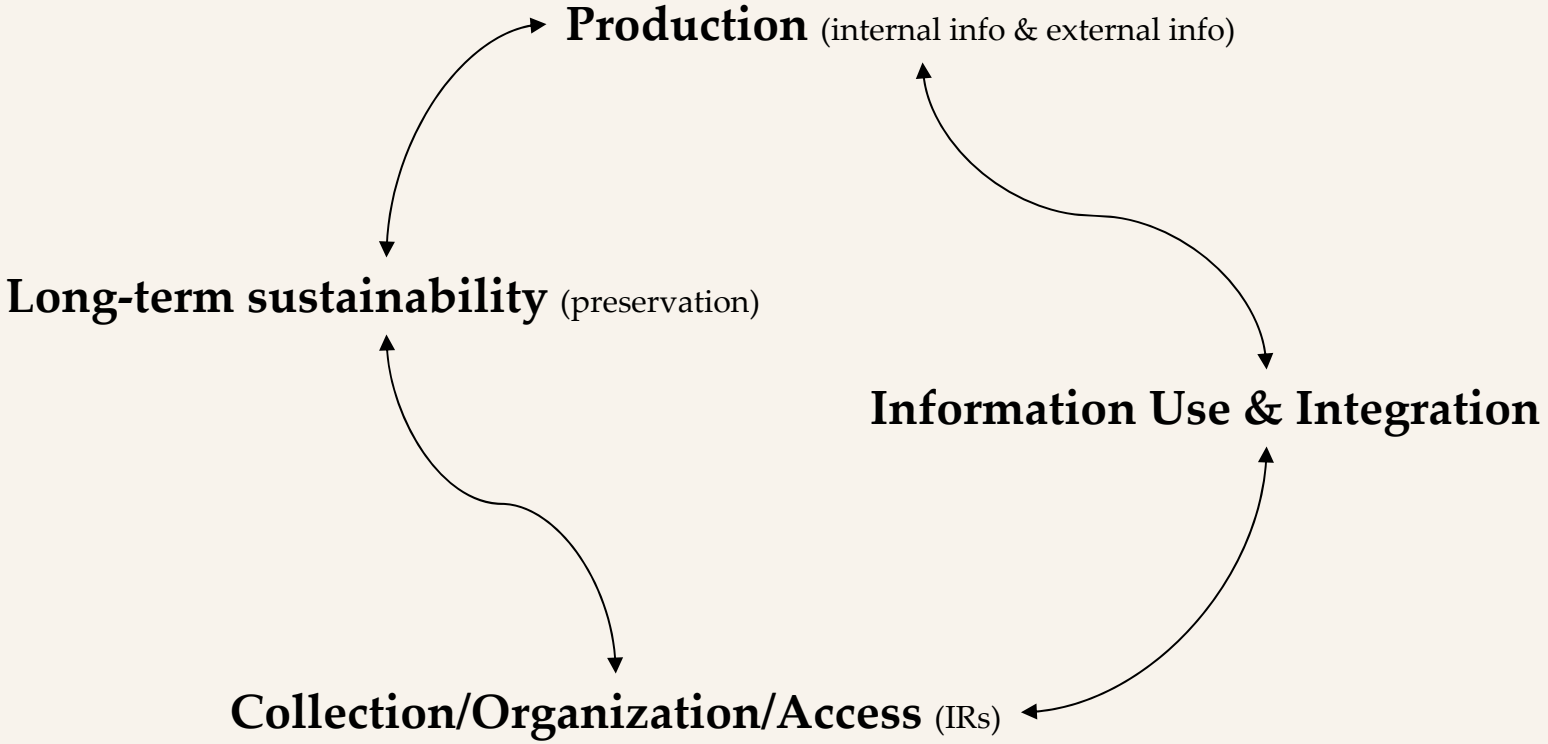
Faculty and Researchers' Scholarly Communications

- **Formal, i.e.**
  - **Journal publications, research papers, technical reports, working papers, conference papers, lectures, records, personal papers**
- **Informal, i.e.**
  - **Listservs, threaded discussion lists, chat, virtual community sites / collaboration spaces, blogs, wikis, e-mail, etc.**
- Conversational / transactional elements of research process more important as libraries capture disciplinary debate and development
- IR-builders should study informal modes, design solutions for capturing and providing additional access to these resources

IR can be a central tool in the challenge of organizing and accessing both formal and informal scholarly communications. Increasingly, will be created, transmitted, and maintained in myriad digital forms

Georgia Tech Library and Information Center

The Life-Cycle of Information

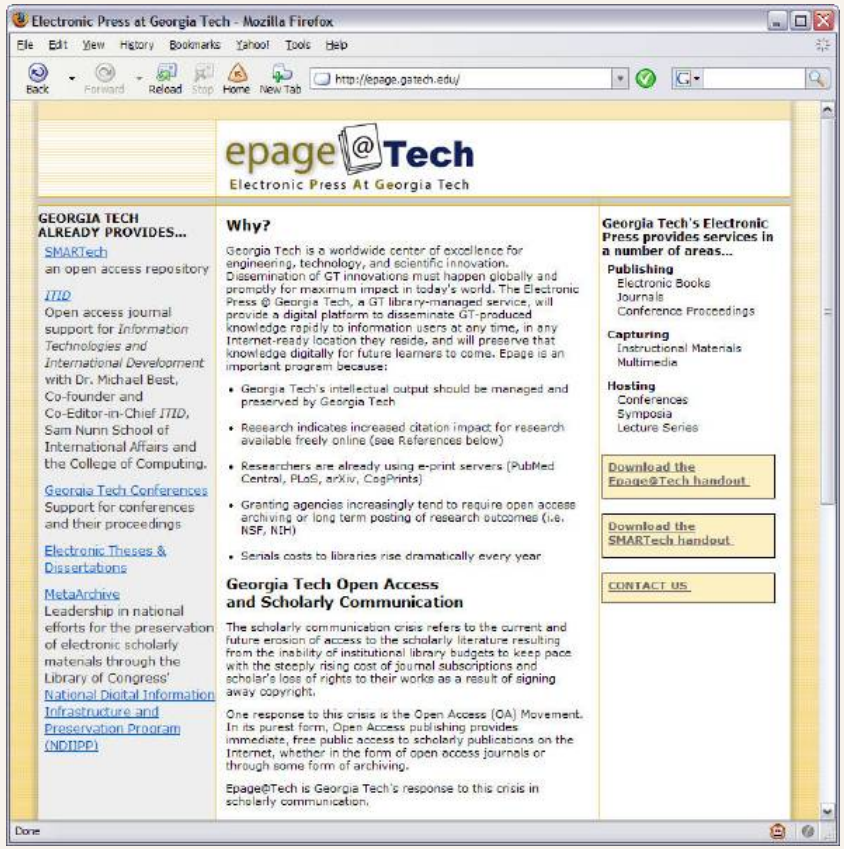


Georgia Tech Library and Information Center

Scholarly Communications Services

- New Library Services:
  - Support creation, use of digital resources in new/different ways

- **Publishing**
  - **Electronic Books**
  - **Journals**
  - **Conference Proceedings**
- **Capturing**
  - **Instructional Materials**
  - **Multimedia**
- **Hosting**
  - **Conferences**
  - **Symposia**
  - **Lecture Series**



Cultivating long-term, active partnerships with faculty is how libraries will continue to transform into high-value hubs of information services

Georgia Tech Library and Information Center

Student Intellectual Output

Undergraduate Research Programs

- Growth of UG research - programs, scholarships, awards, int'l research
- GT: Submit output to SMARTech (via ETD-db software)
  - **GT Examples**:
    - Summer Undergraduate Research Experience
    - Undergraduate Research Scholars Program
    - Undergraduate Research Award

- *GT Journal of Undergraduate Research* (electronic)

* Collaborations with Director of Undergraduate Research, LCC

GT Library East Commons: Student creative activity, research / digital renderings: SMARTech

The Technique (student paper) w/ GT Student Publication Bd, OIT

Georgia Tech Library and Information Center

Learning Objects and other Multimedia Works

LOs proliferating greatly / one of most challenging types of digital output to manage in a repository

What are LOs and why should libraries care?

- Main idea… break educational content down into small chunks that can be reused in various learning environments" (David Wiley)

1) Represent digital intellectual output of univ's faculty, instructors
2) Used in newer, technology-enhanced forms of teaching and learning
3) Librarians will find multimedia objects embedded or linked to more traditional forms of scholarly communications

Librarians need to learn how to manage these objects because they will increasingly find themselves involved with them

Georgia Tech Library and Information Center

Media/Learning Object Repositories

- Outgrowth of student multimedia projects, faculty's learning objects and digitally-captured courses

  GT Examples:

- LWC Multimedia Center / Student output increasingly multimedia, professors ask for it

- Migration from WebCT CE to Sakai – leverage expanded benefits of new virtual Collaboration and Learning Environment
  - Integration of library technologies, content, services

- ECE Digital Media Lab, Distance Learning, Library:
  - Infrastructure-building

Georgia Tech Library and Information Center

Digital Research Data Sets

- New class of digital-born output: digital data sets generated from modern research processes

- **Major Disciplines:**
  - Sciences and engineering fields
  - Social sciences
  - Medical disciplines

- **Examples:**
  - Geospatial data
  - Social science / economic statistical, observational data
  - Biological data
  - Astronomical data
  - Nuclear physics data
  - Genomic and protein data

Georgia Tech Library and Information Center

Digital Research Data Sets

Library Perspective:

- …is a resource to be managed if it is a primary source that must be made available to support and advance research
- …an extension of scholarly publications, e.g., raw, digital data accompanying journal articles and technical papers

Challenges to Libraries:

- Lack of clear policies
- Establish programs, incentives that promote digital data curation as integral part of research projects (i.e. NIH, NSF)
- A need for data curation tools and interoperable technologies
  - tools for data / metadata extraction, database emulation,
    data provenance tracking
- A need for changes: organizational culture, technical infrastructure

Georgia Tech Library and Information Center

Cyberinfrastructure for Managing Intellectual Content

Complexity and diversity of digital intellectual content - librarians, archivists, scholars, and researchers - need a common cyberinfrastructure to manage and foster its utilization

National Science Foundation
Office of Cyberinfrastructure

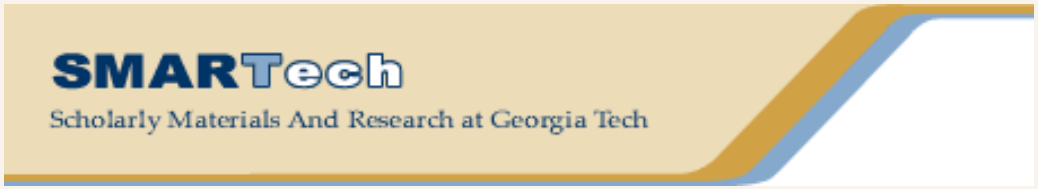
Report of the
American Council
of Learned Societies
Commission on Cyberinfrastructure for
the Humanities and Social Sciences

- **"*Cyberinfrastructure* emphasizes not only technology, but policy and people to form a well integrated whole, fostering scholars' and researchers' participation in communities that advance knowledge and learning" (NSF Cyberinfrastructure report, 2003)**

- **GT has worked on two key cyberinfrastructure components:**
  - **1) Repositories (organization, access, disseminate)**
  - **2) Digital Preservation Networks**

Georgia Tech Library and Information Center

Cyberinfrastructure: Organization Access, Dissemination



"SMARTech, or **S**cholarly **M**aterials **A**nd **R**esearch @ Georgia **Tech**, is a repository for the capture of the intellectual output of the Institute in support of its teaching and research missions"

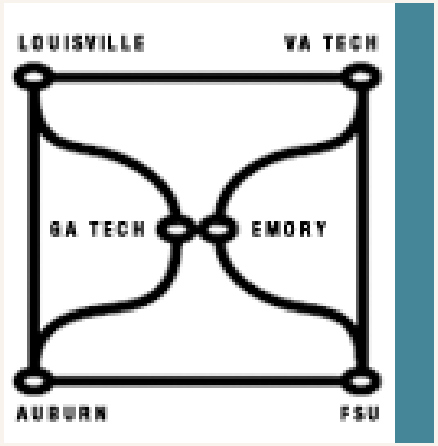
- DSpace installation, opened August, 2004
- 10,000 objects / 70+ (sub)communities / 160+ collections
- 1,000,791 item records viewed, 489,292 items downloaded, and 50,434 searches made (July '05 – June'06)
- SMARTech is the:
  - **4th largest** of 58 DSpace repositories in the United States
  - **13th largest** of 179 DSpace repositories in the world
  - **13th largest** of 201 repositories on any software in U.S.
  - **55th largest** of 764 repositories on any platform in the world

Georgia Tech Library and Information Center

Cyberinfrastructure: Dig. Preservation

- **MetaArchive Preservation Network**
  http://metaarchive.org
  http://smartech.gatech.edu:8282/dspace/handle/1853/7347

  - **Decentralized Approach** (question the "one copy at one institution" approach)
  - **Built on LOCKSS** (supports "distributed digital replication" approach). Peer-to-peer network architecture, each node communicates with each other to replicate content. Minimum of six nodes
  - **Closed Archive** (preservation with no public access. High accessibility = high costs)
  - **Automated format migration tools**
  - **Low Cost** (Planned minimal expense, low barriers to adoption for mid-size insts.)
  - **Flexible, adaptable multi-inst. model**
  - **LC / NDIIPP partnership** (1 of 8 initial)



Georgia Tech Library and Information Center

Thank you!

Tyler O. Walters

Associate Director for Technology & Resource Services

Georgia Tech Library and Information Center

404-385-4489

Tyler@gatech.edu

Georgia Tech Library and Information Center